We thank you and the other reviewer for your insightful review and very constructive comments, which helped improve our paper greatly. We really appreciate your review very much! We also thank other three colleagues who also made very useful comments. We have revised our paper based on the comments from you, the other reviewer and other three colleagues. We also have reworded/rephrased some sentences which we believe can improve the paper. Our responses are listed in below after each reviewers' comment.

**Anonymous Referee #1**

The authors examine dust and smoke aerosols in two different regions in the Atlantic. Situations with opaque water clouds are used to provide a directly retrieved value of the aerosol optical depth. The retrieved aerosol properties are presented and compared with other literature estimates. Furthermore, the constrained optical depths are used to evaluate CALIPSO L2 products. Overall the study presents unique observations of dust/smoke over cloud layers and provides a detailed analyses of CALIPSO L2 products; therefore this reviewer finds that this paper may be suitable for publication after the following concerns are addressed.


MAJOR COMMENTS

pg 13-14 section 3.5: I have some concerns that the FC method may be very sensitive to calibration biases. For example, if the measured backscattered profile was biased high, then even after subtracting the molecular backscatter, the remaining bias in each range bin would be attributed to an aerosol optical depth. Although it is noted that an improved calibration is used (pg 15 lines 18-20), the authors could quantify this possibility by applying the FC method to those profiles which have no aerosol detected in the L2 retrievals (i.e. the subset used in Fig. 3). Those optical depths will be quite noisy, but should be randomly distributed about zero if no calibrations biases exist.

Response: We agree that the FC retrieval is sensitive to the calibration biases (as well as to other error sources) when the aerosol layer is weakly scattering in a relative sense. It is also true for the OWC constrained retrieval as we pointed out in both the original and revised manuscript. We tried to quantify the impact of the possible calibration biases. As an example the results for the dust region are presented in the plot below. Shown in (a) are the averaged attenuated backscatter ratio of all 5km profiles used to determine the clear-air $\gamma'_{WC,SS,NA}$ along with the retrieved extinction profiles suing a lidar ratio (Sa) of 40 and 45 [sr]. There seems to be some residual aerosol and/or calibration bias. The calibration bias seems to be no larger than 2%, looking at the attenuated backscatter ratio of 1.02 around 8 km, which is consistent with the comparison study with the Langley airborne HSRL (please refer to Rogers et al. 2011 ACP paper http://www.atmos-chem-phys.net/11/1295/2011/acp-11-1295-2011.html). Shown in (b) are the histograms of AODs retrieved from each 5km profile using Sa = 40 and 45 [sr]. The mean values are smaller than 0.01. We also attempted to conduct the same analysis using the very recently released V4 L1 data products and made comparisons of the V3 and V4 results. The calibration of the V4 L1 data has been improved greatly. The comparisons show that the mean/median lidar ratio derived for the dust region can be reduced by ~ 1 sr and increased by 1-2 sr in the smoke region. This provides some idea about the possible impact of the calibration biases and residual aerosols.
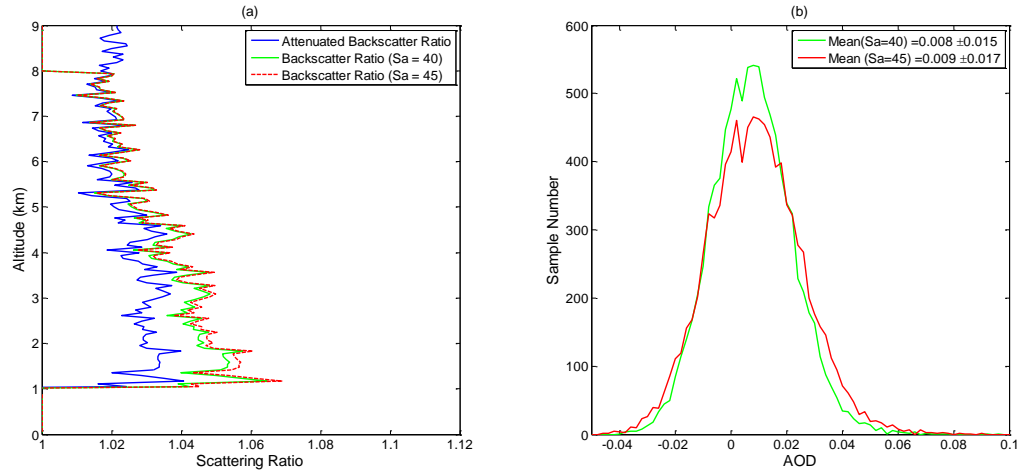
The manuscript would benefit greatly if some effort was made to improve the quality of the figures: many lines are hard to see and distinguish from others and some color scales are poorly designed making it difficult to see patterns. Specifically:

The red boxes in Figs. 6 and 7 I assume correspond to those in Fig 1. Longitude and latitude labels should be added to Fig. 1 to make this clear. Also in-text references to the "red box" confused me at first, I think it would be better to reference these areas as the "spatial domain".

Response: Revised Fig. 1 and its caption. We also used "spatial domain" to replace "geographic region" throughout the paper.

Fig. 2: why does the fit not exist for smaller sizes? The fit appears not to be used anywhere in the manuscript, so I would suggest just removing it entirely.

Response: Removed the fit curve.

Fig. 3: the range on the color scale in panels (b) and (c) should be reduced to highlight the patterns in the red boxes.

Response: Reduced. In addition, we also added one panel of the fraction of calibration clouds (i.e., ratio of number of calibration clouds to the total sample number), as well as panels $S_{wc}$, $H$, and $\gamma'_{WC,MS,NA}$.

Fig. 5: it is hard to see the difference between the dark and light green in panel (c), suggest changing one of them to a different color.

Response: Changed the dark green to brown.

Fig. 6 and 7: reduce the range of the color scales so the patterns in the red boxes can be seen. Also, since the focus of the study is on the domains in the red boxes, I don't think it is necessary to show the regions outside those areas on these plots.

Response: Reduced the color scales. We trimmed out largely the regions surrounding the spatial domain selected and left small part of the surrounding regions in the plots so that readers can have a better idea about the measured distribution and its connection with the surrounding regions.

Figs. 11-13: the red and blue reference lines are very hard to see on the panels that plot

2D histograms. I would suggest making these lines thicker and picking different colors than red or blue since both of those colors are already used to color the histogram.

Response: We have made the lines thicker. We tried other colors and they do not seem better, so we use the current colors in the revised plots.

Figs. 3a, 6a, 7a: instead of plotting the number samples here, it would be nice to see this expressed as a fraction of the total number of CALIPSO samples. This would help provide the reader with the context of how often the OWC technique can be used.

Response: Yes, agree. We added the fraction plots to them and kept the sample number plots.

MINOR COMMENTS

pg 4 line 10: need citation for sentence ending in "year round"

Response: Cited one paper by D. Liu et al. 2008.

pg 4 line 11: need citation for sentence ending in "during the summer"

Response: Cited one paper by D. Liu et al. 2008.

pg 6 eq 1: It doesn't appear that this solution is actually used in the study, so I would suggest removing it from the manuscript.

Response: Removed.

pg 7 section 3.2: much of the information in the section is repeating material already presented in section 2. I would suggest only discussing the geographical regions in section 2 and moving the details of the CALIPSO algorithms to this section.

Response: Thank you for the good suggestion! We have made changes as suggested.

pg 7 line 20: here the authors state that multiple scattering is usually negligible, but for the focus of this study, dust and smoke, that isn't quite true (for example see Wandinger et al. Geophys. Res. Lett. 2010). The authors themselves contradict this statement on pg 25 lines 1-8 where a discussion of potential dust multiple scattering factors are discussed. It should be made clear here that while the CALIPSO retrievals assume no multiple scattering, physically that isn't always true for the aerosols being examined in this study.

Response: The multiple scattering factor $\eta$ is set to one in the current version of the data products. Multiple scattering effects become significant in dense dust (e.g., Liu et al., JQSRT, 2011). However, our recent analyses of multiple scattering effects based on SAMUM measurements and Monte Carlo physical optics modeling indicate multiple scattering for moderate dust is probably at most a 10% effect. We have not been able to duplicate the results of Wandinger et al. (2010), even for large particles. This is supported by comparison with the LaRC airborne HSRL measurement presented in a new figure (Fig. 9) added in the revised manuscript. This comparison needs to be analyzed in more detail, to determine whether or not to use a lower value than $\eta = 1$.

We made revision on page 8, "In the V3 aerosol retrieval, $\eta = 1$ for all aerosol species. Because of this treatment for the multiple scattering, the extinction coefficients and AODs reported in the V3 data products should be considered as effective values, with the multiple scattering contributions that depend to some extent on both the aerosol loading and the aerosol optical and microphysical properties. However, our simulations have shown that multiple scattering is a small effect within moderately dense dust layers and insignificant for smoke (Winker 2003; Liu et al., 2011). Figure 9 supports the idea that multiple scattering effects in moderate dust are small. $\eta$ for dust may be reduced somewhat in the next data release, based on these previous results. Based on comparisons to field measurements and simulations using large particle sizes measured in the source regions, Wandinger et al. (2010) assert that a more significant multiple scattering correction is needed for the CALIOP dust measurements, and that failing to make this correction

can cause underestimates of 10–40% in the extinction retrievals. Further work is planned in this area."

pg 9 line 2: for the re-scaling only a single aerosol type is used so "dominant" should be changed to "only"

Response: Changed to "assuming that only dust or smoke is the dominant aerosol type in the respective region"

pg 10 lines 17-18: The sentence "This ratio is..." is confusing. There is no depolarization term in Eq. (4).

Response: It should be Eq. (5) in the original manuscript, but we deleted this sentence in the revised manuscript because it is not very necessary.

pg 11 line 17: is there a particular justification to using cloud tops lower than 2 km? Seems like the only concern here is to make sure that ice clouds aren't used. But in the tropics the melting layer remains somewhat constant at about 4-5 km, so 2 km seems overly restrictive. Also, why not use the temperature to ensure liquid phase instead of an altitude cutoff?

Response: Because the variability in the water cloud microphysical properties is a major source of the errors in the OWC constrained retrieval, we want to over restrict the selection of water clouds (ideally single type), with a hope that this can help reduce the possible impact of the variability in water cloud physical properties. The temperature is not used because it is not necessary. As mentioned in the manuscript, the top temperature of the selected clouds is generally higher than 8°C.

pg 12 lines 8-21: I found this paragraph to be not clear. The way the lead in is to this paragraph gives the impression that some estimate of particle size is used here. After reading it over a few times I think what is actually done is that the measured backscatter (B prime) is integrated over the water cloud layers and Eq. 5 is used to quantify the multiple scattering contribution. Then the variability of this single-scattering integrated backscatter is attributed to differences in the liquid lidar ratio and hence particle sizes. If that is the case, it might be interesting to see what the spatial distribution of Eq. 5 looks like, since it is possible that the spatial variability presented in Fig. 3 is some combination of difference in multiple scattering and particle size.

Response: We made changes to the paragraph to make it clearer.

The relationship between multiple scattering and depolarization ratio has been well established and studied for water clouds. However, it is a great idea to take a look at the spatial distribution of Eq. 4 (i.e., Eq. 5 in the original manuscript). Shown in below are maps of H and $\gamma'_{WC,SS,NA}$ and $\gamma'_{WC,MS,NA}$ for the opaque water clouds without overlying aerosols and clouds. As seen in this figure, these parameters have a different spatial distribution pattern. There may be some overcorrection or under correction of the multiple scattering due to noise in each profile in $\gamma'_{WC,SS,NA}$, however, it does not seem the spatial pattern in $\gamma'_{WC,SS,NA}$ is due to the multiple scattering. We have added H and $\gamma'_{WC,MS,NA}$ in the revised Figure 3.

(a) Smoothed H  (b) Smoothed $\gamma'_{WC,SS}$  (c) Smoothed $\gamma'_{WC,MS}$
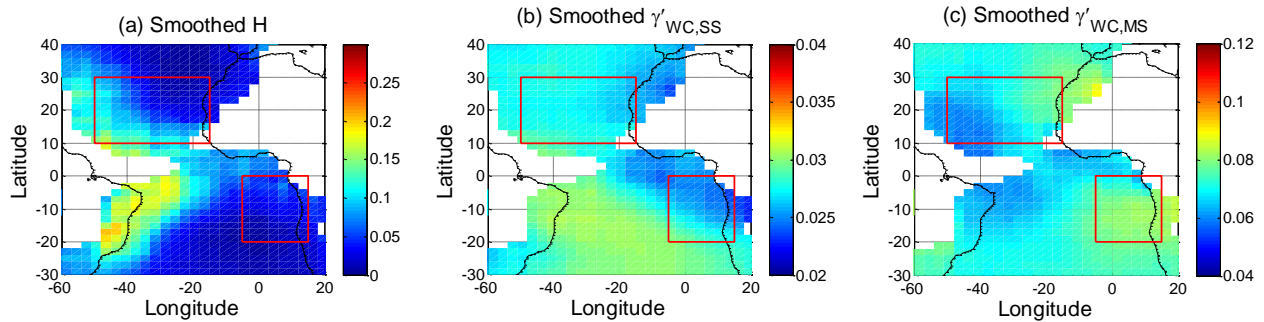
pg 13-14 section 3.5: It is not clear what lidar equation solution is used, is it equation 2?

Yes.

pg 15 lines 1-6: how is the solar background, which is a random error, responsible for a bias? Also is it not correct to say that more averaging can reduce a bias error, only a random error can be reduced by averaging. I think what the authors are getting at is that the solar background necessitates more averaging, which then increases the chances of including broken clouds in the sample. So it's not the solar background causing the biases, as stated in these lines, but the inhomogeneity of the clouds.

Response: We have changed this part to "The analyses were restricted to nighttime measurements, as the large amount of solar background noise present in daytime measurements require signal averaging over longer distances (e.g., 20 km), which in turn would require opaque clouds with corresponding larger horizontal extents and hence significantly reduce the total number of samples available"

pg 15 lines 7-13: most of this can be removed, it is already discussed in section 2

Response: Totally agree. Thank you!

pg 15 lines 22-23: why must the water clouds be single layered? As long total column liquid optical depth is opaque it shouldn't matter how many layers there are.

Response: As I mentioned in the response earlier, we want to over restrict the selection of clouds to reduce some uncertainty due to the variability of the clouds.

pg 17 lines 19-21: the sentence "Sa is an..." is a little too strong of a statement; there are many HSRL and Raman lidar studies that show plenty of variability for a single aerosol type, even within the same profile (i.e. the vertical variation). The authors themselves say this on pg 25 lines 25-27, noting the natural variability within a single aerosol type to be quite large.

Response: Deleted the sentence.

pg 21 lines 3-6: citation(s) needed for these statements

Response: Added three citations.

pg 21 line 17: quantify "well correlated", i.e. give the correlation coefficient

Response: The correlation coefficient is not a very good measure in this context and deleted it in the revised manuscript. Instead we give the slop of the fit curve. Changed the sentence to "For the dust transport region, as shown in Fig. 11a, the majority of $AOD_{L2}$-$AOD_{OWC}$ scatters falls on a line with a slop of ~0.75 (the fit curve is not shown)." on page 22.

pg 22 line 20: quantify "well correlated", i.e. give the correlation coefficient

Response: Please see the response above. Modified the first sentence to "In the smoke transport region, the L2 AOD retrieval is not as successful as in the dust transport region. There are two branches in the AODL2-AODOWC distribution (Fig. 12a)." on page 23

pg 23-25 section 4.5: The discussion of the relationship of lidar ratio error to optical depth error is nice to see. Does this also depend on the type of solution used? For instance I would expect a different relationship between the two for Eqs. 1 and 2 since propagating error through these two equations will give different expressions. If that is the case, it would be nice to point that out in this section.

Response: The closed form (i.e., Eq.(1) in the original manuscript and is currently deleted in the revise manuscript) and iterative form (i.e., Eq.(1) in the revised manuscript) of the solution should be equivalent. We think, the sensitivity of the two solutions to the error should also be similar over the parameter space considered in this paper. However, we will not discuss this in the revised paper because the closed-form solution has been deleted. Young et al (2013) have performed an intensive error sensitivity analysis. More rigorous analyses can be found in that paper. We note that, however, Eq. (8) was derived with an assumption that the aerosol layer is dense or moderately dense so that the molecular scattering is insignificant. We added in the revised manuscript "We note that while Eq. (8) was originally derived under assumption that the aerosol layer is dense or moderately dense, it appears to be equally applicable throughout the whole parameter space considered in this paper. A more rigorous analysis of extinction error propagation and parameter sensitivities can be found in Young et al. (2013)." on page 25.

pg 25 line 16: suggest changing "water clouds," to "water cloud during nighttime,"

Response: Done.

TECHNICAL CORRECTIONS

pg 9 line 2: remove the extra period

Response: Done.

pg 9 line 13: change "(based" to "based"

Response: Done.

pg 10 line 3: remove "recently"

Response: Done.

pg 14 line 9: change "ad" to "and"

Response: Done.

pg 18 lines 15-16: this sentence can be removed

Response: Done.

pg 22 lines 22-23: remove the sentence "The following factors..."

Response: Done.

pg 27 line 8-9: remove underlining

Response: Done.